Magnitude of daily energy deficit predicts frequency but not severity of menstrual disturbances associated with exercise and caloric restriction

Nancy I. Williams¹, Heather J. Leidy¹, Brenna R. Hill¹, Jay L. Lieberman¹, Richard S. Legro², Mary Jane De Souza¹

¹ Women’s Health and Exercise Laboratory and the Department of Kinesiology, Pennsylvania State University, University Park, PA 16802
² M.S. Hershey Medical Center, Department of Obstetrics and Gynecology, Penn State University College of Medicine, Hershey, PA 17033

Abbreviated title: Energy deficit predicts frequency of menstrual disturbances

Corresponding Author:
Nancy I. Williams, ScD
Department of Kinesiology
Women’s Health and Exercise Laboratory
108 Noll Laboratory
Penn State University
University Park, PA 16802
Phone: 814-865-9014
Fax: 814-865-1275
Email: niw1@psu.edu

Author Contributions: NI Williams designed and conducted the study, HJ Leidy assisted in conducting the study, BR Hill aided in data analysis and manuscript writing, JL Lieberman aided in data analysis and manuscript writing, RS Legro and MJ De Souza consulted on menstrual cycle analyses and MJ De Souza assisted in editing.
ABSTRACT

We assessed the impact of energy deficiency on menstrual function using controlled feeding and supervised exercise over four menstrual cycles (1 Baseline and 3 Intervention cycles) in untrained, eumenorrheic women ages 18-30 yrs. Subjects were randomized to either an exercising control (EXCON) or one of three exercising energy deficit (ED) groups i.e., mild (ED1) (-8 ± 2%), moderate (ED2) (-22 ± 3%), or severe (ED3) (-42 ± 3%). Menstrual cycle length and changes in urinary concentrations of estrone-1-glucuronide, pregnanediol glucuronide, and mid-cycle luteinizing hormone were assessed. Thirty-four subjects completed the study. Weight loss occurred in ED1 (- 3.8 ± 0.2 kg), ED2 (- 2.8 ± 0.6 kg), and ED3 (- 2.6 ± 1.1 kg), but was minimal in EXCON (-0.9 ± 0.7 kg). The overall sum of disturbances (luteal phase defects, anovulation, and oligomenorrhea) was greater in ED2 compared to EXCON and greater in ED3 compared to EXCON AND ED1. The average percent energy deficit was the main predictor of the frequency of menstrual disturbances (f = 10.1; β= -0.48; R²=0.23; p=0.003) even when including weight loss in the model. The estimates of the magnitude of energy deficiency associated with menstrual disturbances ranged from -22% (ED2) to -42% (ED3), reflecting an energy deficit of -470 to -810 kcal per day, respectively. This is the first study to demonstrate a dose response relation between the magnitude of energy deficiency and the frequency of exercise-related menstrual disturbances; however, the severity of menstrual disturbances was not dependent on the magnitude of energy deficiency.

Key words: energy balance, menstrual cycle disturbances, luteal phase, amenorrhea, oligomenorrhea
Introduction

Menstrual irregularities in physically active women and female athletes are frequently observed. Commonly reported irregularities and corresponding prevalence rates include luteal phase defects (29%), anovulation (20%), oligomenorrhea (7%), and amenorrhea (37%) (12). A large body of evidence in a variety of mammalian species has demonstrated a causal link between chronic energy deficiency and the suppression of reproductive function involving the central inhibition of gonadotropin releasing hormone (GnRH) pulsatility (34). In humans, long term energy deficiency can result in functional hypothalamic amenorrhea (FHA) and therefore, decrease estrogen exposure, diminishing estrogen’s impact on bone, reproductive, and cardiovascular regulation, often resulting in osteopenia (29, 34), stress fractures (1, 3, 5), transient infertility, dyslipidemia, and impaired endothelial function (14, 15, 26). Prospective studies support a causal role of low energy availability, on the suppression of reproductive function that commonly occurs in physically active women and female athletes (6, 20, 39).

Although the importance of energy availability for the maintenance of reproductive function has been established (34), many important translational issues have not been addressed. Menstrual cycle disturbances, secondary to an energy deficit, occur along a continuum of severity, ranging from subtle disturbances such as luteal phase defects (short or inadequate luteal phases) to anovulation, oligomenorrhea (menstrual cycle length of 36 - 90 days) and the most severe disturbance, amenorrhea (absence of menses for >90 days) (6, 12). Although short term experiments (i.e. 5 days), which reduce energy availability below a threshold of 30 kcal/kg/LBM with the combination of diet and exercise have demonstrated suppressed luteinizing hormone (LH) pulsatility [10], as well as a host of other metabolic alterations (i.e., decreased IGF-1, insulin, leptin concentrations) [10], these studies do not address energy deficiency as it relates to disruptions in follicular development, ovulation, or luteal function, which may occur when exposure to an energy deficit occurs over a single menstrual cycle or multiple cycles. Furthermore, they do not address the time frame over which such disturbances develop, or the association between a given level of energy deficiency and the severity of menstrual disturbances. In the seminal study by Bullen et al. (6) menstrual disturbances were induced by an abrupt onset of exercise and were exacerbated by weight loss, however the magnitude of energy deficiency was not quantified. To address these issues we conducted a randomized prospective experiment to assess the impact of
varying levels of energy deficiency, created through a combination of caloric restriction and exercise, on
menstrual function in young, premenopausal, sedentary women. We hypothesized that there would be a
dose-response relationship between the induction of menstrual disturbances (luteal phase defects,
anovulation, and oligomenorrhea) and the magnitude of energy deficiency such that the intervention
groups experiencing a greater energy deficit would incur a significantly greater incidence of menstrual
cycle disturbances. As well, we hypothesized that the intervention groups experiencing a greater energy
deficit would incur a greater incidence of more severe menstrual cycle disturbances.

Materials and Methods

Experimental Design Overview
This study utilized a randomized prospective design that employed controlled feeding and
supervised laboratory-based exercise over the course of three menstrual cycles in young, untrained,
premenopausal, eumenorrheic women. The study was conducted over three years, with yearly cohorts
recruited in the fall of the academic year and followed until the end of spring semester. The controlled
feeding and exercise training began after the Screening and Baseline periods, each period lasting one
menstrual cycle. All phases of the intervention were anchored to subjects’ menstrual cycles, and each
study phase consisted of one menstrual cycle (Intervention Cycle 1, Intervention Cycle 2, Intervention
Cycle 3). A post study period of one week where diet and exercise remained controlled allowed for post-
intervention measurements. The study design is illustrated in Table 1. Group assignments were based
on varying levels of energy deficiency created through a combination of caloric restriction and exercise
such that one group remained in energy balance and four groups were in different degrees of an energy
deficit. Repeated assessments of menstrual status, metabolic status, and body composition were
conducted.

Subjects
Inclusion criteria for the study included: 1) no history of serious medical conditions, 2) no current
evidence of disordered eating or history of an eating disorder, 3) age 18 – 30 years; 4) weight 45 – 75 kg,
5) body fat 15 – 35% 6) BMI 18 – 25 kg/m², 7) non-smoking, 8) no medication use that would alter
metabolic hormone levels 9) no significant weight loss/gain (± 2.3 kg) in the last year, 10) less than 1
hour of purposeful aerobic exercise/week for the past six months 11) not taking exogenous hormonal
contraceptives for the past 6 months, and 12) documentation of at least two ovulatory menstrual cycles
during screening. Each subject was informed of the purpose, procedures, and potential risks of
participation in the study before signing an informed consent approved by the Penn State University
Biomedical Institutional Review Board. Subjects were recruited in three yearly cohorts using fliers,
newspaper, and radio ads which targeted the University community in the Fall of the academic year.

Screening and Baseline

Screening procedures assessed medical history, menstrual history (based on self-report from the
prior 12 months), current and past physical activity (16), eating attitudes and behaviors, anthropometrics,
physical health and psychological status. Screening and Baseline periods were initiated on day 1 of
subjects' menstrual cycles and lasted from 2 – 3 menstrual cycles. Screening blood samples were
obtained via venipuncture and were analyzed for a complete blood count (CBC), basic chemistry panel,
and an endocrine screen to rule out hormonal or metabolic disease. Each subject underwent a physical
examination by a General Clinical Research Center (GCRC) clinician, and was instructed how to
complete a 3 day diet log by a GCRC dietician. An interview to rule out current clinical eating disorders,
assess the risk of developing disordered eating behaviors, and to rule out existing major axis I psychiatric
disorders was performed under the direction of a clinical psychologist. Menstrual status was determined
using menstrual calendars, ovulation detection kits (First Response (Tambrands, Inc.) and three mid-
luteal serum progesterone measures (at least one > 5.0 ng/ml) during the Baseline cycle to confirm
ovulation. Menstrual calendars were used to document menstrual bleeding (and therefore cycle length),
and the occurrence of cramps or other menstrual symptoms.

If subjects were deemed to be in good health and had at least two successive ovulatory cycles
they were randomized into subject groups.

Subject Groupings

During the Baseline period, subjects were randomly assigned to an experimental group for the
Intervention Cycles 1, 2, and 3 of the study. The goal of the subject groupings was to test the impact of
varying levels of an energy deficit created by the combination of caloric restriction and exercise on
menstrual function. They were assigned to either a control group that did not exercise and consumed an
amount of calories estimated to maintain body weight, a control group that exercised, but received extra
food calories to remain in energy balance (exercising controls or EXCON), or one of four groups that
exercised and were prescribed reduced energy intake to create varying levels of an energy deficit (energy
deficit or ED groups). ED groups were defined by an energy prescription comprised from the quantity of
calories provided as food and the quantity of calories expended as exercise. ED groups were prescribed
targeted reductions in energy intake (7 days/week) compared to their Baseline energy needs ranging
from –15% to –30% in combination with prescribed increases in exercise energy expenditure (5
days/week) equivalent in calories to +15% to +30% of Baseline energy needs. Specifically, the initial
four energy deficit groups were intended to represent 1) an increase of 15% kcals of exercise (15%
deficit), 2) an increase of 30% kcals of exercise (30% deficit), 3) a decrease of 15% in dietary intake,
combined with an increase of 15% of exercise, (30% deficit) and 4) a decrease of 30% in dietary intake,
combined with an increase of 30% kcals of exercise (60% deficit).

This resulted in actual calculated daily energy deficits ranging from –9 to –42% of Baseline energy needs calculated as follows:

Reduction in food intake (kcal) = % reduction in dietary intake/100 * Baseline energy needs
Exercise energy expenditure (kcal) = % increase in exercise energy expenditure/100 * Baseline energy needs
Total energy deficit (kcal) = reduction in food intake (kcal) + exercise energy expenditure (kcal)
Energy deficit (%) = 100 - (Total energy deficit/Baseline energy needs*100)

Data for the two groups intended to have a 30% deficit were combined in the present analyses to
represent a single energy deficit group, as the actual daily energy deficits calculated for these two groups
were not significantly different (i.e., -20±8% and -28±10%). Data for the control group that did not
exercise are not presented in the current analyses, as the most appropriate comparison group to test the
stated hypotheses is the EXCON group. Therefore, data are presented from the EXCON, and three
energy deficit groups, i.e., ED1, ED2, and ED3 that exhibited increasing average energy deficits of –9.5 ±
1.8%, –22.6 ± 2.5%, and –42.4 ± 7.9% (p < 0.05; ED1 vs. ED2, ED2 vs. ED3, and ED1 vs. ED3). The
exercise and diet intervention lasted the equivalent of three menstrual cycles (Intervention Cycles 1 – 3)
after Baseline for each subject, with the beginning of the intervention falling on the first day of menses
(Intervention Cycle 1).
Dropouts

One hundred and nine subjects met all screening criteria and began the Baseline phase of the study. Forty-seven subjects were excluded during the Baseline period, mostly because they experienced subtle menstrual disturbances, i.e., luteal phase defects and or short or long cycle lengths. Sixty two subjects were randomized into experimental groups including the control group that did not exercise (data not reported). Forty subjects completed the study in the EXCON and ED1-3 groups, and 13 in the control group that did not exercise (not included in present analysis). The dropout rate during the intervention was 14.5%. Dropouts differed significantly in some descriptive characteristics, i.e., they were older, (22 ± 1 vs. 20 ± 0.3 yrs; p<0.05), taller, (170 ± 3 vs. 164 ± 1 cm; p<0.05), and had a higher percentage of body fat, (28.3± 0.8 % vs. 24.2 ± 1.5, p<0.05). Six additional subjects’ data were not included in the current analyses because they exhibited a luteal phase disturbance or an oligomenorrheic cycle (36 days or >) during their Baseline cycle which could only be determined after the study was completed and all urinary assays and detailed analyses of menstrual calendars and menstrual disturbances were performed. The subjects who did not complete intervention or were eliminated for menstrual disturbances in the Baseline cycle were, for EXCON, 1 dropout, and 1 eliminated from analyses; for ED1, 3 dropouts and 3 eliminated from analyses; for ED2, 4 dropouts and 1 eliminated from analyses; and for ED3, 1 dropout and 1 eliminated from analyses. Thus the final n for analysis was 34 subjects in groups: EXCON (n=8), ED1 (n=6), ED2 (n=12), and ED3 (n=8).

Assessment of Baseline Energy Needs

We assessed Baseline energy needs during the Baseline cycle. Twenty-four hour energy expenditure was assumed to represent Baseline energy needs, as all subjects were weight stable (± 1.5 kg) during Screening. To determine 24 hour energy expenditure, resting metabolic rate (RMR) (kcal/24hr) and calories attributable to non-exercise physical activity throughout the day were summed. RMR was measured once during days 1 – 7 of the Baseline cycle using indirect calorimetry using previously reported methods (17). Subjects wore a research accelerometer (RT3, Stayhealthy, Monrovia, CA) for 24 hours per day (except during training bouts, swimming, and bathing), during the first 7 days of the follicular phase of the Baseline cycle for the assessment of non-exercise physical activity (PA) caloric
The calories from the RMR and the non-exercise PA calories were added together. This sum was operationally defined as the Baseline energy needs calorie level (17). This calorie level was verified and adjusted if body weight varied during a one week “calibration” period during Baseline (see below).

**Determination of Caloric Intake during the Intervention**

Depending on the group assignment (EXCON, ED1, ED2, or ED3) the prescribed dietary intake (Intake; kcal), for the intervention was calculated as a percentage reduction from the estimate of Baseline energy needs as determined at the end of the calibration period (described below). All food consumed during the study was prepared and weighed to the nearest gram to achieve the desired calorie level by GCRC metabolic kitchen staff. Energy content of prepared food items were verified with bomb calorimetry by mixing and grinding each day’s menu items and then burning the food for that day in the calorimeter. This was repeated for each of the days’ food from the eight day menu rotation. Subjects were required to eat two of three meals per day during the week at the GCRC dining room. Dinner meals, a daily snack and weekend meals were packed out. Intake was comprised of 55% carbohydrates, 30% fat, and 15% protein. The dietary protocol utilized an eight-day meal rotation. All subjects were also prescribed a multivitamin. The subjects were fed for 7 days during Baseline, and ± 100 kcal adjustments in Intake were made if weight changed ± 0.5 kg at the middle or end of the calibration week. To meet the target level of Intake during the Intervention, food intake was either increased or decreased according to group assignment. Subjects were instructed to eat all and only the food provided to them by the study. Any uneaten food was re-weighed and recorded for later subtraction from the Intake total. Eating food not provided by the study was highly discouraged, but if this occurred, extra food was recorded on a log sheet. Intake and macronutrient composition were calculated using Nutritionist Pro (First Data Bank, Indianapolis, IN).

**Estimation of Energy Balance during the Intervention**

To estimate energy balance during the intervention, measures of Intake, RMR, non-exercise PA, and exercise training energy expenditure in kcal (Ex) were repeated during the intervention and these data were used to calculate daily values for energy balance. The calories for RMR, non-exercise PA, and Ex were summed to represent 24 hour energy expenditure. RMR was assessed during the follicular
phase of Baseline, Intervention Cycle 2, and during the Post-Study period. Non-exercise PA was calculated every other week from subjects wearing the RT3 for seven days, 24 hours per day (except during training bouts, swimming, and bathing). Ex was measured during each training bout (see below).

Intake was calculated daily (see below). Energy balance was thus calculated daily as \( \text{Intake (kcal)} - (\text{RMR} + \text{non-exercise PA} + \text{Ex}) \text{ (kcal)} \) using the most recent data available. Calculations for energy balance were updated continually throughout the intervention as new measurements were obtained.

Metabolic Hormones

Fasting blood samples were obtained during the Baseline, Intervention Cycle 2, and during the Post-Study period for the measurement of triiodothyronine (T3) (ng/dl) and insulin-like growth factor-1 (IGF-1)(ng/ml). The purpose of these measurements was provide an additional laboratory measurement of metabolic status.

Exercise training

All exercise training took place in Noll laboratory and was supervised by trainers. Exercising groups performed aerobic exercise 5 times/week at 70 – 80% of maximum heart rate as determined from tests of maximal aerobic capacity (VO\(_2\)\(_{\text{max}}\)). The duration of exercise (minutes) was equal to that required to expend a prescribed number of calories and ranged from 20 minutes to 75 min depending on the prescription. The total amount of calories expended during each exercise session was measured using the OwnCal feature on the Polar S610 heart rate monitor (Polar Electro Oy, Kempele, Finland)(10).

The heart rate monitors were continually reinitialized twice weekly with the most recent values of weight, maximum heart rate, VO\(_2\)\(_{\text{max}}\), and age. Modes of exercise included treadmill running, elliptical machine exercise, stair stepping, and stationary cycling. When subjects came to the training room, they removed their RT3 monitors (used to monitor non-exercise PA kcal) and were given their Polar heart rate monitor. When they finished their workout, they removed the Polar and put the RT3 back on.

Menstrual Status

Menstrual cycle length and menstrual symptoms were documented throughout the study with menstrual calendars. Daily urine samples were collected during the Baseline and Intervention Cycles 1-3 according to previously published methods (36). Estrogen and progesterone exposure, confirmation of
ovulatory status, the presence or absence of luteal phase defects, and the lengths of the follicular and luteal phases were determined by analysis of daily urinary metabolites of estrone-1-glucuronide (E1G), pregnanediol glucuronide (PdG), and mid-cycle luteinizing hormone (LH) after completion of the study using previously published methods (11). Menstrual disturbances characterized included luteal phase defects, oligomenorrheic cycles, and anovulatory cycles. Two types of luteal phase defects were assessed, short luteal phases (< 10 days in length) and inadequate luteal phases (PdG peak concentration < 5.0 µg/ml (11, 32)). Oligomenorrheic cycles were defined as cycles 36 days or greater in length (12). Anovulatory cycles were defined as cycles lacking an adequate pre-ovulatory E1G peak, an absence of a mid-cycle LH surge, and no luteal rise in PdG above 2.49 µg/ml (11, 32). We quantified menstrual cycle disturbances two ways. For each subject, the sum of the total number of menstrual disturbances experienced over the course of three menstrual cycles was calculated. For example, a subject who experienced an anovulatory cycle in Intervention Cycle 2 and an oligomenorrheic cycle that also had a short luteal phase in Intervention Cycle 3 would have a sum of total disturbances of three. We also tabulated whether each subject experienced a menstrual disturbance or not in each of the Intervention Cycles 1-3. This allowed us to determine the number of subjects in each experimental group who experienced at least one occurrence of each type of menstrual disturbance. We expressed this as the # of subjects with a given disturbance/total number of subjects, expressed as a ratio and as a percentage.

Body Composition and Body Weight Measurements

Body weight was measured to the nearest 0.01 kg using a digital scale (Seca; Hamburg, Germany) twice per week in the morning before eating breakfast by the GCRC dietician while subjects wore a standard shorts and tee shirt. During Baseline, once during Intervention 2 or 3, and once during the Post-Study period subjects’ body composition was determined once during the follicular phase (days 1 – 7) using underwater weighing as previously described (17).

Resting Metabolic Rate

RMR was measured between 0600 – 1000 hours following an overnight fast. Upon arrival at the laboratory, the subject lay in a supine position on a bed for 20 – 30 minutes to acclimate to room temperature and undergo familiarization with the equipment and procedures. A ventilated hood was then
placed over the subject’s head for 30 minutes and metabolic rate was measured using previously published methods (17).

**Aerobic Capacity**

Maximal aerobic capacity ($VO_2\text{max}$) was determined during Screening, Baseline, each Intervention Cycle, and during Post-Study. Maximal aerobic capacity ($VO_2\text{max}$) was determined from an incremental exercise test on a treadmill to volitional exhaustion using the Modified Astrand Protocol using indirect calorimetry (Sensormedics Metabolic Cart Model no. 229, Sensormedics Corp., Yorba Linda, CA)(28).

**Blood Sampling**

For all blood samples, subjects were fasted overnight prior to blood collection which occurred between 0700 and 1000 hours at the GCRC after lying supine for at least 15 min.. Blood was allowed to clot for one hour prior to centrifugation. Serum was stored at -20°C until later analysis.

**Biochemical Analyses**

CBC/Chem 24 and the endocrine screen determinations were completed by Quest Diagnostics (Lyndhurst, NJ). The endocrine screen included LH, follicle stimulating hormone (FSH), prolactin, progesterone, estradiol, thyroxine and thyroid-stimulating hormone. Circulating progesterone was measured using a radioimmunoassay (RIA) on serum samples (Diagnostics Products Corporation, Los Angeles, CA). The sensitivity of this assay is 0.02ng/mL. The intra- and inter-assay coefficients of variation for the high controls were 2.7% and 3.9%, respectively. The intra- and inter-assay coefficients of variation for the low controls were 8.8% and 9.7%, respectively. Competitive enzyme immunoassays (EIA) were used to measure the major urinary estrogen metabolite, estrone-1-glucuronide (E1G) and pregnanediol glucuronide (PdG) according to previously published methods (23). Urinary LH was measured using RIA (Diagnostic Products Corporation, Los Angeles, CA). The sensitivity of this assay is 0.15 mIU/mL. The intra- and inter-assay coefficients of variation for the high controls were 1.0% and 3.4%, respectively. The intra- and inter-assay coefficients of variation for the low controls were 1.6% and 7.1%, respectively. Circulating total $T_3$ was measured using a radioimmunoassay for total $T_3$ (Diagnostics Products Corporation, Los Angeles, CA). The sensitivity of this assay is 7 ng/dl. The intra-assay and inter-assay coefficients of variation for the high controls were 3.1% and 5.7%, respectively; the intra-
Intraclass correlation coefficients of variation for the low controls were 2.4% and 5.2%. All biochemical determinations were assayed in duplicate and all samples from a given subject were measured in the same assay.

Statistics and Data Analysis

All data sets were tested for non-normality and outliers before statistical hypothesis tests were performed. Outliers detected were rejected and thus, excluded from analysis. Data analyzed included those from all subjects who completed the study. Analyses for normally distributed data were conducted using general linear model procedures for repeated measures with four groups to test for changes over time, group differences, and time X group interactions. Post hoc analyses to isolate main effects were conducted using the Bonferroni correction. Analyses for the frequency of menstrual disturbances, i.e., the sum of total disturbances and the sum for each type of disturbance, were performed using a one way ANOVA. Post hoc testing to isolate main effects was performed using t-tests with least significant difference comparing each group to the EXCON group. To determine differences between groups in the occurrence of menstrual disturbances in each cycle of the Intervention, Chi Square analyses were conducted. To determine predictors of menstrual disturbances when all subjects are considered together, stepwise linear regression analyses were performed. Data were analyzed using SPSS for Windows (version 18.0; Chicago, IL). All data are expressed as mean ± SEM and p < 0.05 was considered statistically significant.

Results

Descriptive Characteristics

Descriptive data from all subjects are illustrated in Table 2. There were no significant differences among ED groups in parameters of age, anthropometrics, initial VO\textsubscript{2}max, reproductive maturity, menstrual cycle length, initial fitness, or caloric intake.
Compliance to Diet and Exercise Prescriptions

Prescribed caloric intake and actual caloric intake varied less than 35 kcal per week during the intervention, and no differences in this parameter were found among groups. With respect to exercise training, 95% of the subjects consistently reached their prescribed exercise kcal expenditure per week and target exercise intensity.

Energy Balance Parameters

Table 3 depicts energy balance parameters averaged across the intervention for each group. Dietary intake was significantly lower in ED1 and ED2 vs. EXCON, and significantly lower in ED3 vs. all other groups. Twenty-four hour energy expenditure, non-exercise physical activity throughout the day, and RMR were not different among groups. Exercise kcal per workout session was significantly lower in the ED1 group vs. all other groups. By design, the calculated energy balance (deficit/surplus) was significantly different when each group was compared to the other groups, whether calculated as kcal/day or as % of baseline energy needs. RMR did not change significantly over time (P=0.643), and did not differ among groups (P=0.757). There was not a significant group X time interaction (P=0.093). Non-exercise PA did not change over time (P=0.953), and there were no differences among groups (P=0.212). There was not a significant group X time interaction (P=0.691).

Effects of the Intervention on Maximal Aerobic Capacity, Body Weight, and Body Composition, and Metabolic Hormones

The actual training intensity throughout the intervention was not different among groups and was 76-79% of maximal heart rate (range 73%-89%) for all subjects. Maximal aerobic capacity improved significantly (Time effect p < 0.05) by 18% above Baseline values, but the change did not differ among groups (all subjects at baseline VO\textsubscript{2max} = 37.9 ± 0.7 ml/kg/min and post VO\textsubscript{2max} = 44.2 ± 1.6 ml/kg/min).

Body weight declined significantly over time (p<0.001) and there was a significant time X group interaction (p = 0.005). Weight loss occurred in ED1 (- 3.8 ± 0.2 kg), ED2 (- 2.8 ± 0.6 kg), and ED3 (- 2.6 ± 1.1 kg), whereas EXCON experienced minimal weight fluctuation (-0.9 ± 0.7 kg). (Figure 1). Figure 1 depicts the change from Baseline in body composition parameters. Fat loss accounted for most of the change in weight (time effect p < 0.0001) since no significant changes in fat free mass were observed. Percent fat declined significantly (time effect p < 0.001) mostly attributable to the ED1, ED2 and ED3...
groups (time X group p = 0.002). There were no differences between groups in Baseline T3; EXCON
92.7 ± 6.5 ng/dl, ED1 116.2 ± 10.2 ng/dl, ED2 111.3 ± 4.9 ng/dl, and ED3 107.3 ± 2.7 ng/dl (P>0.05), but
T3 declined significantly in all groups except EXCON. Similarly, IGF-1 was not different between groups
at Baseline; EXCON 379.4 ± 30  mg/ml, ED1 346.8 ± 23 mg/ml, ED2 399.9 ± 30 mg/ml, and ED3 366.0
±29 mg/ml, but declined significantly over time in all groups except EXCON.

Effects of the Intervention on Menstrual Cycle Characteristics

No significant changes in menstrual cycle length or follicular phase length were observed in any
group (Table 4). Luteal phase length declined significantly over time in all groups (p < 0.0001) but on
average, only achieved the criteria for “short” in the ED3 group. When changes over time were examined
in each group luteal phase length decreased in the ED3 group (time effect p =0.01) at Intervention Cycle
3. (p<0.001). Table 5 shows the percentage and numbers of subjects in each group who experienced
each type of menstrual disturbance during each intervention cycle. The most frequent clinically significant
menstrual disturbance observed was the development of luteal phase defects, i.e., short luteal phases
and inadequate luteal phases. The frequency of these disturbances, which were mostly short luteal
phases, increased as the intervention progressed from 20% to 38% of all subjects experiencing luteal
defects from Intervention Cycle 1 to Intervention Cycle 3. The frequency of luteal phase disturbances
increased significantly from EXCON to ED1 to ED2 and ED3, with the lowest percentage of subjects in
EXCON (8%) and the highest percentage of subjects in ED3 (52%) experiencing luteal defects in at least
one intervention cycle (p = 0.053). Alternatively, the frequency of oligomenorrheic and anovulatory cycles
did not exhibit a dose response relationship across increasing severity of energy deficiency, as there
were no differences between groups. Figure 2 exhibits the incidence of each menstrual disturbance for
the entire intervention in each group. The incidence of luteal phase defects was greater in ED2 and ED3
when compared to EXCON. ED3 also incurred a significantly greater number of luteal phase defects
when compared to ED1 (panel A). Neither the incidence of annovulatory (panel B) nor oligomenorrheic
(panel C) cycles is significantly different among groups. Figure 3 shows the relation between the
magnitude of energy deficit experienced by each group as the intervention progressed (panel A) and the
overall frequency of all types of menstrual disturbances (panel B). When compared to EXCON both the
ED2 and ED3 groups incurred a significantly greater number of total disturbances. ED3 also incurred a significantly greater number of total disturbances when compared to ED1.

When all subjects from all groups were considered together in a stepwise multiple linear regression analysis to predict the frequency of all menstrual disturbances, a dose response relationship was observed between the frequency of all types of disturbances vs. the magnitude of the energy deficit. That is, the average percent energy deficit was a significant predictor of the frequency of menstrual disturbances ($F = 10.1; \beta = -0.49; R^2 = 0.24; p = 0.003$). Interestingly, weight loss in kg, fat loss in kg, and the change in RMR were not significant predictors of the frequency of menstrual disturbances when included in the model.

**Discussion**

This is the first study to demonstrate a dose response relationship between the magnitude of energy deficiency and the frequency of exercise-related menstrual disturbances. As well, this study demonstrates two important clinical findings: 1) When young women experience the equivalent of three menstrual cycles’ exposure to an energy deficit ranging from -8 to -42% of their typical energy needs, the most frequent form of menstrual disturbance that occurs is luteal phase defects characterized by either a short (<10 days) or inadequate (progesterone < 5.0 µg/ml) luteal phase, and 2) the proportion of subjects experiencing a luteal phase defect and the frequency of luteal phase defects were greater in groups experiencing a greater energy deficit (ED2 and ED3) i.e., moderate or severe, when compared to the control group and the ED1, or mild energy deficit group. The average sum of luteal phase defects resulting from the moderate and severe deficits exceeded prevalence rates previously reported in exercising women (2, 11, 12). When all subjects and all menstrual disturbances are considered together our regression analyses indicated that the magnitude of energy deficit is linearly related to the overall frequency of menstrual disturbances.

Suppression of reproductive function, i.e., FHA, in response to chronic energy deficiency may occur as a protective mechanism to preserve fuel for life-sustaining processes in the body (35). The neuroendocrine mechanism through which this suppression occurs is unclear but involves a complex interplay whereby peripheral signaling of metabolic status to hypothalamic GnRH neurons results in a decrease in GnRH pulsatile activity and a suppression of the reproductive axis (22). The goal of the
current study was to address a more practical question involving the magnitude of energy deficiency that
causes these effects. To this end, we conducted a randomized, prospective study to assess the impact of
caloric restriction and exercise designed to produce mild, moderate, and severe levels of energy
deficiency on menstrual function in young previously sedentary women.

Our subjects were typical of college age women in terms of body weight, composition, and
fitness. By design, they did not engage in purposeful exercise training or dieting prior to the study, and
they exhibited very “robust” menstrual function, i.e., regular, ovulatory menstrual cycles. Effects of the
exercise and caloric restriction intervention could then be ascribed to the treatment versus any
predisposition to disrupted menstrual function. Our intervention produced modest weight loss, and typical
losses in body fat percentage and gains in aerobic fitness. This intervention could be characterized as
what a previously untrained woman would experience if she underwent a rather vigorous exercise
program for a short period of time.

The estimates of the magnitude of energy deficiency associated with menstrual disturbances can
be estimated from the moderate (ED2) and severe (ED3) energy deficit groups and thus ranged from –
22% to – 42% compared to baseline energy needs, reflecting an energy deficit of – 470 to – 810 kcal per
day. This magnitude of an energy deficit was substantial enough to result in clinical and subclinical
menstrual disturbances, which, if experienced over a long time frame, could result in other clinical
sequelae associated with the female athlete triad (24). It is notable that this change resulted from both
increases in energy expenditure associated with exercise training combined with decreases in food
intake. Our finding that the magnitude of energy deficiency is directly related to the induction of menstrual
disturbances extends previous laboratory findings from a short term prospective study by Loucks et al.
(20). The latter study demonstrated a suppression of LH pulsatility after five days below an energy
availability of 30 kcal/kg LBM (18) and in the current study we are showing that a longer term exposure to
energy deficiency causes subclinical menstrual cycle disturbances in a dose dependent manner. Notably,
the most common disturbance observed, i.e., shortening of the luteal phase length, occurred with no
significant change in menstrual cycle length or follicular phase length. Thus, no apparent or physical
indication of a menstrual disturbance occurred and subtle menstrual disturbances were detected
exclusively by the measurement of urinary metabolites of menstrual cycle hormones. Consequently,
exercising women may be experiencing subtle forms of menstrual dysfunction secondary to an energy
deficit, but may not be aware of such disturbances or the potential impact of these disturbances on fertility
and bone health (13, 18, 25).

Several participants also presented with oligomenorrheic and/or anovulatory menstrual cycles in
response to the diet and exercise intervention. The frequency of these more severe menstrual
disturbances was not statistically significant among groups, and was fairly similar to previous reports in
exercising women or in a large study in women aged 18-44 years (11, 12, 33). The lack of difference
between groups may be physiologically significant in that an energy deficit in some participants elicited
more severe menstrual disturbances than the same energy deficit in other participants. Thus, individuals
may vary greatly in their sensitivity to changes in energy balance and susceptibility to particular types of
menstrual dysfunction (27).

The current study also extends the prior analyses of the seminal study by Bullen et al. (6), who
performed an intervention in a similar population to determine the impact of an abruptly imposed exercise
program, with and without weight loss, on menstrual function. The previous study was conducted in 28
college-aged, untrained women over the course of two menstrual cycles. Participants were required to
run four miles/day, progressing to ten miles/day by the fifth week and engage in 3½ hours of moderate-
intensity sport activity. Two groups, a weight maintenance group and weight loss group, were utilized to
determine the impact of exercise, with or without weight loss, on menstrual function. The study
demonstrated that 24 of 28 subjects experienced either luteal phase defects (short and or inadequate
luteal phases) or suppression of the pre-ovulatory LH surge over the course of the study. Specifically,
delayed menses, luteal phase defects and loss of the LH surge occurred in both groups, but in a
significantly higher proportion of subjects in the weight loss group (weight loss up to 75% of subjects vs.
weight maintenance up to 44% of subjects), who lost on average 4.0 ± 0.3 kg, when compared to the
weight maintenance group who lost 1.0 ± 0.2 kg. By comparison, our findings were similar to this study in
terms of the percentage of luteal phase defects produced (2) but fewer subjects in our study experienced
anovulatory cycles (moderate 29% and severe 42% vs. up to 81%) than in Bullen et al. (6). That study
utilized body weight as a primary outcome measure and thus, was not designed to measure changes in
the components of energy balance. The weight stable exercising control group lost only 1.0 ± 0.2 kg body
weight but still exhibited menstrual disturbances and thus, may have been exposed to an energy deficit
with no statistically significant change in body weight.

Menstrual dysfunction in the absence of significant changes in body weight has also been
demonstrated in monkeys where amenorrhea was observed in response to an energy deficit in the
absence of significant weight loss (38). As well, a lack of changes observed in body weight in response
to subtle changes in energy balance may be a result of expansion of plasma volume (9), increases in
muscle mass from training, and/or increased body water stored with training-induced increases in
glycogen storage all of which can mask decreases in body weight that might result from negative energy
balance. Thus, the use of body weight as a determinant of changes in energy balance with exercise
training may not truly reflect subtle changes in energy status and therefore be misleading when used as
an index of metabolic stress on the reproductive axis. Our intervention produced a significant change in
body weight similar to the weight loss group in the study by Bullen et al. but the changes in body weight
were not predictive of menstrual disturbances while estimates of the daily energy deficit were.

In other studies, varying levels of energy deficiency have been shown to elicit inconsistent
responses with regard to suppression of reproductive function. This may be due to differences in
susceptibility to reproductive dysfunction in response to energy balance changes and may be a
consequence of gynecological age as a more mature reproductive axis exhibits less susceptibility to
disruption (19, 30, 40). Other factors such as stress responsiveness (4, 7) may play a role in individual
differences in susceptibility. There is evidence that stressors can act synergistically, which would add
variability in reproductive responses to a metabolic stressor (37). Recent evidence of a genetic basis for
FHA has also emerged demonstrating that there may be gene mutations that exist in the gene encoding
the GnRH receptor that may increase susceptibility of women to exercise-associated menstrual
disturbances (8). Despite evidence that other factors may indeed be at play when considering one’s
susceptibility to exercise associated menstrual disturbances, the dose response effect of energy
deficiency on menstrual disturbances we observed in the current study underscores the strength of
energy availability as a strong modulator of menstrual function.

This study has many strengths. It is a prospective study which provides an extension of prior,
cross-sectional studies (6) and is thus, more generalizable to women that may engage in regular
exercise. Caloric intake was carefully controlled and all exercise was supervised for the duration of the study. Additionally, the measurement of daily urinary metabolites of menstrual cycle hormones allowed for the determination of subtle menstrual disturbances like LPD which would otherwise have gone unnoticed had other methods of documenting menstrual status such as calendars or sporadic blood sampling been used.

A limitation of the study is a small sample size. A greater number of participants experienced oligomenorrheic cycles in the groups experiencing a greater energy deficit; however, this proportion was not significantly different between groups. Thus, we may not have had enough power to demonstrate that a significantly greater proportion of participants were experiencing more severe menstrual disturbances, i.e. oligomenorrhea in response to a greater energy deficit. However, studies such as the current one are difficult to perform as they are costly and require extreme subject compliance and thus, subject burden. We utilized carefully controlled feeding and supervised exercise to create a quantifiable energy deficit across three menstrual cycles and have thus, been able to present accurate and quantifiable data with regard to suppression of menstrual function in response to changes in energy balance. Moreover, the time course of the study may not have allowed for more severe disturbances, i.e. oligo- and amenorrhea, to be observed in response to the energy deficit. However, this study is still, to date, the longest prospective intervention, other than those conducted in an animal model, designed to assess the magnitude and severity of menstrual disturbances that occur in response to varying levels of energy deficiency. It is possible that if the intervention were carried out longer, more frequent and more severe disturbances might have developed. The idea that subtle disturbances might be preceded more severe disturbances in supported by Bullen et al. (6) and by a study in exercising female monkeys where menstrual cycles immediately before induction of amenorrhea were characterized by suppressed progesterone concentrations and lower early follicular luteinizing hormone (LH) concentrations (38, 39).

Thus, the subtle menstrual disturbances observed in the current study may increase the risk of, or subsequently lead to more severe menstrual disturbances. Another limitation may be the generalizability of the results. We included only subjects in whom two ovulatory cycles were documented during screening and in whom Baseline cycles were ovulatory and devoid of other defects such as LPD. Thus, the subjects were very homogenous with respect to menstrual status. While this approach is an important
feature of the experimental design in order to draw conclusions about an intervention, our results may not be generalizable to the many young women that experience irregular menstrual cycles for unidentified reasons. Given the strict screening criteria we used (those with history of menstrual disturbances or recent disturbances excluded, and two consecutive ovulatory menstrual cycles documented prior to intervention) it is likely that our subjects were robust in terms of their susceptibility to our intervention. It is likely that a similar intervention conducted in subjects with a wider range of baseline menstrual function would have had a greater disruptive effect.

In conclusion, a combination of caloric restriction and exercise that produced varying levels of energy deficiency ranging from -470 to -810 kcal per day over a period of three menstrual cycles induced menstrual disturbances which occurred more frequently in women who were the most energy deprived. Because individuals may vary greatly in their sensitivity to changes in energy balance and susceptibility to menstrual dysfunction, more research is needed to determine how this effect is modified by age, training status, initial menstrual function, susceptibility to psychological stress, and other factors. Moreover, additional research on the effects of similar interventions on other endpoints such as bone structure and strength are warranted.

Acknowledgments

We thank Ann Cathcart, Thom Parrott, Brian Frye, Kelly Dougherty, Meredith Snook, Erica Richards, and Jackie Gardner for their important contributions to this research. We also appreciate the extraordinary cooperation of the study subjects, and the expert assistance of the GCRC staff.

Grants

This study was supported by NIH RO1 HD39245-01 (N.I.Williams) and NIH M01RR10732.

Department of Kinesiology, Women’s Health and Exercise Laboratory, Penn State University.

Disclosures

The authors have nothing to disclose.

References


Figure Legends:

Figure 1. Change in body weight (kg) (A), fat free mass (kg) (B), fat mass (kg) (C) and body fat (%) (D) T3 (E), and IGF-1 (F) from Baseline to Intervention 1 (Int1), Intervention 2 (Int 2) and Intervention 3 (Int 3) (A) or Baseline, Mid-Study, and Post-Study (B-F) in all groups. For A body weight (kg) EXCON n=8, ED1 n=6, ED2 n=12, ED3 n=8; B fat free mass (kg) EXCON n=6, ED1 n=6, ED2 n=11, ED3 n=8; C fat mass (kg) EXCON n=6, ED1 n=6, ED2 n=11, ED3 n=8; D percent fat (%) EXCON n=6, ED1 n=6, ED2 n=11, ED3 n=8; E T3 EXCON n=6, ED1 n=5, ED2 n=11, ED3 n=8; F IGF-1 EXCON n=6, ED1 n=5, ED2 n=11, ED3 n=8; for A, B, C, D, and E ANOVA with repeated measures main effects with post hoc testing using Bonferroni correction to isolate time effect: a= ED1 vs. Baseline p<0.0125; b=ED2 vs. Baseline p<0.0125, c=ED3 vs. Baseline p<0.0125; For panel F, Friedman test for related samples: a= ED1 p<0.05, b=ED2 p<0.05, c=ED3 p<0.05; Data are reported as mean ± SEM

Figure 2. Frequency of each menstrual disturbance observed across the intervention in each group. Data are reported as mean ± SEM of the total sum of disturbances per group. a significantly different from EXCON. b significantly different from EXCON and ED1; Data are reported as mean ± SEM

Figure 3. A. Average daily energy deficit (%) experienced across study phases and B. Sum of all menstrual disturbances across the intervention for each group. Data are reported as mean ± SEM: A: a ED1 significantly different from EXCON. b ED2 significantly different from EXCON. c ED2 significantly different from EXCON and ED1. d ED3 significantly different from EXCON, ED1 and ED2; B: a significantly different from EXCON. b significantly different from EXCON and ED1; Data are reported as mean ± SEM
Figure 2.
Figure 3.
Table 1. Experimental procedures

<table>
<thead>
<tr>
<th>Screening</th>
<th>Baseline Cycle</th>
<th>Intervention Cycle 1</th>
<th>Intervention Cycle 2</th>
<th>Intervention Cycle 3</th>
<th>Post-Study (Cycle Days 1-7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 Day Diet Log</td>
<td>3-Day Diet Log</td>
<td>3-Day Diet Log</td>
<td>3-Day Diet Log</td>
<td>3-Day Diet Log</td>
<td>3-Day Diet Log</td>
</tr>
<tr>
<td>Body Weight</td>
<td>Body Weight</td>
<td>Body Weight</td>
<td>Body Weight</td>
<td>Body Weight</td>
<td>Body Weight</td>
</tr>
<tr>
<td>Menstrual Calendar</td>
<td>Menstrual Calendar</td>
<td>Menstrual Calendar</td>
<td>Menstrual Calendar</td>
<td>Menstrual Calendar</td>
<td>Menstrual Calendar</td>
</tr>
<tr>
<td>Ovulation Detection</td>
<td>Ovulation Detection</td>
<td>Supervised Exercise &amp; Controlled Feeding</td>
<td>Supervised Exercise &amp; Controlled Feeding</td>
<td>Supervised Exercise &amp; Controlled Feeding</td>
<td>Supervised Exercise &amp; Controlled Feeding</td>
</tr>
<tr>
<td>Mid-luteal Serum Progesterone</td>
<td>Mid-luteal Serum Progesterone</td>
<td>VO₂max RMR Activity kcal</td>
<td>VO₂max RMR Activity kcal</td>
<td>VO₂max RMR Activity kcal</td>
<td>VO₂max RMR Activity kcal</td>
</tr>
<tr>
<td>Medical History</td>
<td>Calibration Period (7 days) to determine energy needs</td>
<td>Exercise kcal</td>
<td>Exercise kcal</td>
<td>Exercise kcal</td>
<td>Exercise kcal</td>
</tr>
<tr>
<td>Menstrual History</td>
<td>VO₂max RMR Activity kcal</td>
<td>Randomization to Experimental Groups</td>
<td>VO₂max RMR Activity kcal</td>
<td>VO₂max RMR Activity kcal</td>
<td>VO₂max RMR Activity kcal</td>
</tr>
<tr>
<td>Physical Activity</td>
<td>Exercise kcal</td>
<td></td>
<td>Exercise kcal</td>
<td>Exercise kcal</td>
<td>Exercise kcal</td>
</tr>
<tr>
<td>Eating Attitudes</td>
<td>Exercise kcal</td>
<td></td>
<td>Exercise kcal</td>
<td>Exercise kcal</td>
<td>Exercise kcal</td>
</tr>
<tr>
<td>CBC/Chemistry</td>
<td>Exercise kcal</td>
<td></td>
<td>Exercise kcal</td>
<td>Exercise kcal</td>
<td>Exercise kcal</td>
</tr>
<tr>
<td>VO₂max Panel</td>
<td>Exercise kcal</td>
<td></td>
<td>Exercise kcal</td>
<td>Exercise kcal</td>
<td>Exercise kcal</td>
</tr>
<tr>
<td>Panel</td>
<td>Exercise kcal</td>
<td></td>
<td>Exercise kcal</td>
<td>Exercise kcal</td>
<td>Exercise kcal</td>
</tr>
<tr>
<td>Endocrine Screen</td>
<td>Exercise kcal</td>
<td></td>
<td>Exercise kcal</td>
<td>Exercise kcal</td>
<td>Exercise kcal</td>
</tr>
<tr>
<td>Physical Exam</td>
<td>Exercise kcal</td>
<td></td>
<td>Exercise kcal</td>
<td>Exercise kcal</td>
<td>Exercise kcal</td>
</tr>
<tr>
<td>Psychological Interview</td>
<td>Exercise kcal</td>
<td></td>
<td>Exercise kcal</td>
<td>Exercise kcal</td>
<td>Exercise kcal</td>
</tr>
<tr>
<td>MC</td>
<td>MC</td>
<td>MC</td>
<td>MC</td>
<td>MC</td>
<td>MC</td>
</tr>
<tr>
<td>Menstrual Cycle 1</td>
<td>Menstrual Cycle 2</td>
<td>Menstrual Cycle 3</td>
<td>Menstrual Cycle 4</td>
<td>Menstrual Cycle 5</td>
<td>Menstrual Cycle 6</td>
</tr>
</tbody>
</table>

MC=Menstrual Calendar

Note: Mid-study Body composition testing occurred during Intervention Cycle 2 for most subjects, but in some it occurred during Intervention Cycle 3.
<table>
<thead>
<tr>
<th>Demographics</th>
<th>EXCON (n = 8)</th>
<th>ED1 (n = 6)</th>
<th>ED2 (n = 12)</th>
<th>ED3 (n = 8)</th>
<th>p-value group effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (yr)</td>
<td>20.5 ± 0.4</td>
<td>20.5 ± 0.8</td>
<td>20.2 ± 0.6</td>
<td>20.4 ± 0.7</td>
<td>0.989</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>165 ± 3.0</td>
<td>165 ± 3.0</td>
<td>164 ± 2.0</td>
<td>164 ± 2.0</td>
<td>0.945</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>57.9 ± 2.2</td>
<td>56.4 ± 1.2</td>
<td>59.8 ± 1.2</td>
<td>59.8 ± 1.5</td>
<td>0.385</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>21.4 ± 1.3</td>
<td>20.7 ± 0.7</td>
<td>22.2 ± 0.5</td>
<td>22.4 ± 0.6</td>
<td>0.496</td>
</tr>
<tr>
<td>Body Composition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Body Fat (%)</td>
<td>26.5 ± 2.4</td>
<td>27.4 ± 1.5</td>
<td>28.8 ± 1.0</td>
<td>29.6 ± 1.3</td>
<td>0.528</td>
</tr>
<tr>
<td>Fat Mass (kg)</td>
<td>15.6 ± 2.0</td>
<td>15.5 ± 1.2</td>
<td>17.3 ± 0.8</td>
<td>17.8 ± 1.5</td>
<td>0.511</td>
</tr>
<tr>
<td>Fat Free Mass (kg)</td>
<td>42.3 ± 1.0</td>
<td>40.9 ± 0.6</td>
<td>42.6 ± 1.0</td>
<td>42.1 ± 1.2</td>
<td>0.729</td>
</tr>
<tr>
<td>Reproductive Characteristics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age of Menarche (yr)</td>
<td>12.3 ± 0.7</td>
<td>12.2 ± 0.2</td>
<td>11.9 ± 0.3</td>
<td>12 ± 1.0</td>
<td>0.949</td>
</tr>
<tr>
<td>Baseline Cycle Length (days)</td>
<td>28.6 ± 0.9</td>
<td>27.8 ± 0.6</td>
<td>28.7 ± 0.5</td>
<td>30.2 ± 0.8</td>
<td>0.165</td>
</tr>
<tr>
<td>Training/Diet Characteristics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VO₂ peak (ml/kg·min⁻¹)</td>
<td>36.1 ± 1.4</td>
<td>38.4 ± 2.1</td>
<td>37.6 ± 1.4</td>
<td>38.1 ± 1.7</td>
<td>0.797</td>
</tr>
<tr>
<td>Baseline Caloric Intake (kcal)</td>
<td>1548 ± 195</td>
<td>2150 ± 271</td>
<td>2032 ± 154</td>
<td>1881 ± 177</td>
<td>0.198</td>
</tr>
<tr>
<td>Energy Balance Parameters</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMR (kcal/24 hr)</td>
<td>1205 ± 57</td>
<td>1293 ± 27</td>
<td>1285 ± 64</td>
<td>1242 ± 62</td>
<td>0.732</td>
</tr>
<tr>
<td>Non-Exercise PA (kcal/24 hr)</td>
<td>593 ± 64</td>
<td>576 ± 47</td>
<td>737 ± 68</td>
<td>623 ± 64</td>
<td>0.276</td>
</tr>
<tr>
<td>24hr Energy Expend.(kcal)</td>
<td>1797± 92</td>
<td>1870 ± 53</td>
<td>2021 ± 105</td>
<td>1865 ± 77</td>
<td>0.341</td>
</tr>
</tbody>
</table>

Values are expressed as mean ± SEM; PA= physical activity

afrom 3 day diet logs
TABLE 3. Energy balance parameters averaged across Intervention Cycles 1-3 for each group

<table>
<thead>
<tr>
<th></th>
<th>EXCON (n = 8)</th>
<th>ED1 (n = 6)</th>
<th>ED2 (n = 12)</th>
<th>ED3 (n = 8)</th>
<th>p-value</th>
<th>Group Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dietary Intake (kcal/day)</td>
<td>2236 ± 87</td>
<td>1845 ± 57(^a)</td>
<td>1814 ± 43(^a)</td>
<td>1321 ± 17(^{a,b,c})</td>
<td>&lt;0.0001</td>
<td></td>
</tr>
<tr>
<td>24 Hour Energy Expenditure (kcal/day)</td>
<td>2149 ± 82</td>
<td>2018 ± 71</td>
<td>2330 ± 77</td>
<td>2213 ± 85</td>
<td>0.084</td>
<td></td>
</tr>
<tr>
<td>Exercise (kcal/workout)</td>
<td>536 ± 23</td>
<td>323 ± 16(^a)</td>
<td>495 ± 35(^b)</td>
<td>507 ± 30(^b)</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td>Non-Exercise Physical Activity (kcal/day)</td>
<td>575 ± 49</td>
<td>586 ± 58</td>
<td>703 ± 50</td>
<td>672 ± 60</td>
<td>0.265</td>
<td></td>
</tr>
<tr>
<td>Resting Metabolic Rate (kcal/day)</td>
<td>1256 ± 43</td>
<td>1230 ± 67</td>
<td>1333 ± 41</td>
<td>1250 ± 40</td>
<td>0.288</td>
<td></td>
</tr>
<tr>
<td>Energy Deficit/Surplus (kcal/day)</td>
<td>+80 ± 34</td>
<td>-162 ± 45(^a)</td>
<td>-470 ± 51(^{a,b})</td>
<td>-813 ± 72(^{a,b,c})</td>
<td>&lt;0.0001</td>
<td></td>
</tr>
<tr>
<td>Energy Deficit/Surplus (%)</td>
<td>+4 ± 2</td>
<td>-8 ± 2(^a)</td>
<td>-22 ± 3(^{a,b})</td>
<td>-42 ± 3(^{a,b,c})</td>
<td>&lt;0.0001</td>
<td></td>
</tr>
</tbody>
</table>

Values are expressed as mean ± SEM.

\(^{a}\) significantly different from EXCON; \(p <0.016\) post hoc test using Bonferroni correction

\(^{b}\) significantly different from ED1; \(p <0.016\) post hoc test using Bonferroni correction

\(^{c}\) significantly different from ED2; \(p <0.016\) post hoc test using Bonferroni correction
TABLE 4. Effects of the intervention on menstrual cycle length

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>Baseline Cycle</th>
<th>Range</th>
<th>Intervention Cycle 1</th>
<th>Range</th>
<th>Intervention Cycle 2</th>
<th>Range</th>
<th>Intervention Cycle 3</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>EXCON</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cycle Length (days)</td>
<td>8</td>
<td>28.6 ± 0.9</td>
<td>(25 – 32)</td>
<td>29.6 ± 1.3</td>
<td>(25 – 37)</td>
<td>28.9 ± 0.9</td>
<td>(25 – 32)</td>
<td>27.0 ± 0.6</td>
<td>(24 – 28)</td>
</tr>
<tr>
<td>Follicular Phase (days)</td>
<td>8</td>
<td>16.4 ± 0.6</td>
<td>(14 – 19)</td>
<td>17.5 ± 1.3</td>
<td>(14 – 26)</td>
<td>17.5 ± 1.0</td>
<td>(13 – 21)</td>
<td>15.5 ± 0.6</td>
<td>(13 – 18)</td>
</tr>
<tr>
<td>Luteal Phase (days)</td>
<td>8</td>
<td>12.3 ± 0.6</td>
<td>(10 – 15)</td>
<td>12.1 ± 0.5</td>
<td>(11 – 14)</td>
<td>11.4 ± 0.7</td>
<td>(7 – 13)</td>
<td>11.5 ± 0.5</td>
<td>(10 – 13)</td>
</tr>
<tr>
<td>ED1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cycle Length (days)</td>
<td>6</td>
<td>27.8 ± 0.6</td>
<td>(26 – 30)</td>
<td>29.1 ± 1.1</td>
<td>(26 – 33)</td>
<td>27.3 ± 1.0</td>
<td>(24 – 32)</td>
<td>29.5 ± 1.0</td>
<td>(27 – 34)</td>
</tr>
<tr>
<td>Follicular Phase (days)</td>
<td>6</td>
<td>16.8 ± 0.9</td>
<td>(14 – 20)</td>
<td>17.5 ± 1.0</td>
<td>(15 – 21)</td>
<td>16.0 ± 1.7</td>
<td>(14 – 22)</td>
<td>19.3 ± 1.7</td>
<td>(15 – 25)</td>
</tr>
<tr>
<td>Luteal Phase (days)</td>
<td>6</td>
<td>11.0 ± 0.4</td>
<td>(10 – 13)</td>
<td>11.7 ± 0.8</td>
<td>(10 – 15)</td>
<td>11.3 ± 2.1</td>
<td>(2 – 18)</td>
<td>10.1 ± 1.1</td>
<td>(5 – 12)</td>
</tr>
<tr>
<td>ED2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cycle Length (days)</td>
<td>12</td>
<td>28.7 ± 0.5</td>
<td>(26 – 31)</td>
<td>27.5 ± 1.1</td>
<td>(21 – 35)</td>
<td>27.3 ± 1.2</td>
<td>(22 – 34)</td>
<td>28.5 ± 1.4</td>
<td>(24 – 41)</td>
</tr>
<tr>
<td>Follicular Phase (days)</td>
<td>7</td>
<td>16.0 ± 0.5</td>
<td>(14 – 20)</td>
<td>17.8 ± 0.9</td>
<td>(14 – 22)</td>
<td>17.3 ± 1.0</td>
<td>(14 – 23)</td>
<td>17.6 ± 1.4</td>
<td>(11 – 23)</td>
</tr>
<tr>
<td>Luteal Phase (days)</td>
<td>7</td>
<td>12.4 ± 0.4</td>
<td>(10 – 13)</td>
<td>11.0 ± 1.0</td>
<td>(5 – 14)</td>
<td>10.4 ± 1.2</td>
<td>(5 – 15)</td>
<td>10.6 ± 1.0</td>
<td>(6 – 14)</td>
</tr>
<tr>
<td>ED3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cycle Length (days)</td>
<td>8</td>
<td>29.6 ± 0.6</td>
<td>(28 – 35)</td>
<td>28.4 ± 0.7</td>
<td>(26 – 36)</td>
<td>28.6 ± 1.5</td>
<td>(25 – 59)</td>
<td>30.2 ± 3.1</td>
<td>(22 – 44)</td>
</tr>
<tr>
<td>Follicular Phase (days)</td>
<td>5</td>
<td>18.5 ± 1.3</td>
<td>(13 – 23)</td>
<td>18.2 ± 1.4</td>
<td>(15 – 24)</td>
<td>21.0 ± 2.2</td>
<td>(13 – 30)</td>
<td>20.6 ± 2.7</td>
<td>(14 – 34)</td>
</tr>
<tr>
<td>Luteal Phase (days)</td>
<td>5</td>
<td>11.6 ± 0.9</td>
<td>(10 – 15)</td>
<td>9.4 ± 1.5</td>
<td>(4 – 15)</td>
<td>6.8 ± 2.1</td>
<td>(3 – 15)</td>
<td>6.8 ± 1.2</td>
<td>(3 – 13)</td>
</tr>
</tbody>
</table>

Values are expressed as mean ± SEM;
- time p < 0.001; group p=0.069; group x time p=0.184; ANOVA with repeated measures on all groups, Luteal Phase (days)
- time p = 0.01 vs. Baseline Cycle; ANOVA with repeated measures on each group separately
- p < 0.001 post hoc paired t-test; Intervention Cycle 3 vs. Baseline Cycle

Note: subjects who experienced anovulatory cycles are not included in the analyses for follicular or luteal phase length
<table>
<thead>
<tr>
<th></th>
<th>EXCON</th>
<th>ED1</th>
<th>ED2</th>
<th>ED3</th>
<th>Total All Groups</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Luteal Phase Defects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intervention Cycle 1</td>
<td>0/8</td>
<td>0/6</td>
<td>5/12</td>
<td>2/7</td>
<td>7/33</td>
<td>0.07 (χ²=7.0)</td>
</tr>
<tr>
<td>Percentage (%)</td>
<td>(0)</td>
<td>(0)</td>
<td>(42)</td>
<td>(29)</td>
<td>(21)</td>
<td></td>
</tr>
<tr>
<td>Intervention Cycle 2</td>
<td>1/8</td>
<td>1/6</td>
<td>4/10</td>
<td>4/6</td>
<td>10/30</td>
<td>0.14 (χ² = 5.5)</td>
</tr>
<tr>
<td>Percentage (%)</td>
<td>(13)</td>
<td>(13)</td>
<td>(40)</td>
<td>(67)</td>
<td>(33)</td>
<td></td>
</tr>
<tr>
<td>Intervention Cycle 3</td>
<td>0/8</td>
<td>1/6</td>
<td>4/9</td>
<td>5/6</td>
<td>10/29</td>
<td>0.008* (χ² = 11.8)</td>
</tr>
<tr>
<td>Percentage (%)</td>
<td>(0)</td>
<td>(29)</td>
<td>(44)</td>
<td>(83)</td>
<td>(34)</td>
<td></td>
</tr>
<tr>
<td><strong>Intervention Total</strong></td>
<td>1/8</td>
<td>1/6</td>
<td>10/12</td>
<td>7/8</td>
<td>19/34</td>
<td>0.001* (χ² = 16.8)</td>
</tr>
<tr>
<td>Percentage (%)</td>
<td>(13)</td>
<td>(38)</td>
<td>(83)</td>
<td>(88)</td>
<td>(56)</td>
<td></td>
</tr>
<tr>
<td><strong>Oligomenorrheic Cycles</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intervention Cycle 1</td>
<td>1/8</td>
<td>0/6</td>
<td>0/12</td>
<td>0/8</td>
<td>1/34</td>
<td>0.34 (χ² = 3.3)</td>
</tr>
<tr>
<td>Percentage (%)</td>
<td>(13)</td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
<td>(3)</td>
<td></td>
</tr>
<tr>
<td>Intervention Cycle 2</td>
<td>0/8</td>
<td>0/6</td>
<td>0/12</td>
<td>1/8</td>
<td>1/34</td>
<td>0.34 (χ² = 3.3)</td>
</tr>
<tr>
<td>Percentage (%)</td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
<td>(13)</td>
<td>(3)</td>
<td></td>
</tr>
<tr>
<td>Intervention Cycle 3</td>
<td>0/8</td>
<td>0/6</td>
<td>1/12</td>
<td>2/7</td>
<td>3/33</td>
<td>0.20 (χ² = 4.6)</td>
</tr>
<tr>
<td>Percentage (%)</td>
<td>(0)</td>
<td>(0)</td>
<td>(8)</td>
<td>(29)</td>
<td>(9)</td>
<td></td>
</tr>
<tr>
<td><strong>Intervention Total</strong></td>
<td>1/8</td>
<td>0/6</td>
<td>1/12</td>
<td>2/7</td>
<td>4/33</td>
<td>0.43 (χ² = 2.8)</td>
</tr>
<tr>
<td>Percentage (%)</td>
<td>(13)</td>
<td>(0)</td>
<td>(8)</td>
<td>(29)</td>
<td>(12)</td>
<td></td>
</tr>
<tr>
<td><strong>Anovulatory Cycles</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intervention Cycle 1</td>
<td>0/8</td>
<td>0/6</td>
<td>0/12</td>
<td>1/8</td>
<td>1/34</td>
<td>0.34 (χ² = 3.3)</td>
</tr>
<tr>
<td>Percentage (%)</td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
<td>(13)</td>
<td>(3)</td>
<td></td>
</tr>
<tr>
<td>Intervention Cycle 2</td>
<td>0/8</td>
<td>0/6</td>
<td>2/12</td>
<td>2/8</td>
<td>4/34</td>
<td>0.32 (χ² = 3.5)</td>
</tr>
<tr>
<td>Percentage (%)</td>
<td>(0)</td>
<td>(0)</td>
<td>(17)</td>
<td>(25)</td>
<td>(12)</td>
<td></td>
</tr>
<tr>
<td>Intervention</td>
<td>Subjects with at least one Anov. cycle / # subjects</td>
<td>Percentage (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------------------------------</td>
<td>----------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0/8</td>
<td>1/8</td>
<td>3/12</td>
<td>0/7</td>
<td>4/33</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0)</td>
<td>(13)</td>
<td>(25)</td>
<td>(0)</td>
<td>(12)</td>
<td></td>
</tr>
</tbody>
</table>

$\chi^2 = 5.7$)

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Subjects with at least one Anov. cycle / # subjects</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0/8</td>
<td>0/6</td>
</tr>
<tr>
<td></td>
<td>(0)</td>
<td>(0)</td>
</tr>
</tbody>
</table>

$\chi^2 = 7.0$)

Note: Intervention totals reflect the number of subjects that had at least one menstrual disturbance during Intervention 1-3. If subjects had more than one disturbance they were only counted once. The Total All Groups column reflects the number of subjects who had at least one disturbance when all groups are considered. If subjects had more than one disturbance they were only counted once.

*p < 0.05; LPD = Luteal phase defect; Anov. = anovulatory cycle; Oligo. = oligomenorrheic cycle*